DEVELOPMENT OF A HYBRID HANDS-OFF HUMAN COMPUTER INTERFACE BASED ON ELECTROMYOGRAM SIGNALS AND EYE-GAZE TRACKING

E. C. Lyons, A. B. Barreto, M. Adjouadi
Electrical and Computer Engineering Department
Florida International University, Miami, Florida, USA

Abstract-**A hybrid hands-off human computer interface that uses infrared video eye gaze tracking (EGT) and Electromyogram (EMG) signals is introduced. This system combines the advantages of both sub-systems, providing quick cursor displacement in long excursions and steady, accurate movement in small position adjustments. The hybrid system also provides a reliable clicking mechanism. The evaluation protocol used to test the system is described and the results for the hybrid, an EMG-only interface, and the standard hand-held mouse are described and compared. These results show that the hybrid system is, in an average, faster than the EMG-only system by a factor of 2 or more.**

***Keywords* - Hands-off Computer Interface, Eye Gaze Tracking, Electromyogram, Computer cursor control**

I. INTRODUCTION[1]

Effective interaction with Personal Computers is a basic requirement for many of the functions that we fulfill in our daily lives. Furthermore, the importance of human-computer interaction continues to grow, as more and more of our activities, such as banking, commerce, etc., are transplanted to the Internet. A large proportion of the computers with which we interact every day use a Graphic User Interface (GUI), where the point-and-click paradigm is utilized to select and activate icons. Unfortunately, a significant number of individuals in our society cannot operate the standard "mouse" used to move the screen cursor and perform selections in the GUI, due to severe motor disabilities. This would be the case of quadriplegics, which are unable to use their arms and hands to interact with the computer.

In the last decades several approaches have been attempted to facilitate the interaction between individuals with severe motor disabilities and GUI-driven computers. One of these systems is the infrared-video Eye Gaze Tracking (EGT) system, which is able to estimate, on a real-time basis, the point on the screen where a user is looking, i.e., the estimated "point-of-gaze" (POG). Our experimentation with this kind of systems has confirmed references found in the literature [3, 5, 6] , which indicate the remarkable ability of these systems to displace the cursor across long distances on the computer screen, quickly. However, we have also confirmed the reported inherent instability of the POG estimation and the difficulty of implementing clicking mechanisms with these systems [3,5,6].

Many individuals with severe motor disabilities are unable to move their arms and legs, but retain control over their facial muscles. This observation has prompted our efforts towards the development of a cursor control mechanism activated through voluntary contractions of some cranial muscles. We have developed a prototype that detects the electromyogram (EMG) signals associated with the contraction of the temporalis muscles and the muscles involved in the raising and lowering of the eyebrows, through three electrodes [1,2]. The signals form these electrodes are processed by a Digital Signal Processing (DSP) board and, as a result, the screen cursor of the computer is stepped left, right, up or down. In contrast with the EGT cursor control system, this EMG-based interface is highly stable, keeping the screen cursor static if the user does not perform a voluntary contraction of these muscles. Additionally, we have implemented an intuitive and successful mechanism to perform a selection operation (i.e. clicking) with this interface. To do this, the user only needs to contract both temporalis muscles simultaneously (full jaw clench). However, due to the incremental position modification implicit in the operation of this interface, moving the cursor across long distances in the computer screen may require some time.

Our experiences with both the EGT-based system and our EMG-based interface made us realize that their strengths and weaknesses were complementary and prompted us to develop a "hybrid" interface that incorporates both systems, relying on one or the other for the effective control of the screen cursor, according to the context in which that movement is intended by the user. This paper reports on the design, development and evaluation of such system.

II. METHODOLOGY

Our hybrid EGT/EMG computer interface incorporates both individual interfaces and decides, on a real-time basis, which subsystem should, in fact, govern the movement of the screen cursor. Therefore, this section will give a brief overview of each of the subsystems and then explain how the context of the user input is used to enable one or the other mode of interaction.

A. Eye-Gaze Tracking (EGT) Subsystem

We use an ASL-504 Eye Tracker system by Applied Science Laboratories (ASL) for the implementation of this subsystem. This system follows the same infrared video eye gaze tracking principles pioneered by the ERICA system [4]. The ERICA system was based on a near-infrared light source that illuminates the eyes of the computer user, who sits in front of the computer screen, while a video camera, with an infrared lens, continuously captures images of one or both of the user's eyes. Using this infrared process, reflections at two particular points in the user's eyes can be obtained: the first is the bright reflection of the illumination on the cornea, or "Corneal Reflection" (CR), and the second one is the bigger reflection observed in the pupil, the "Pupil Reflection" (PR).

[1] This work was sponsored by NSF Grants: EIA-9812636 and EIA-9906600

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25 Oct 2001	N/A	-

Title and Subtitle Development of a Hybrid Hands-Off Human Computer Interface Based on Electromyogram Signals and Eye-Gaze Tracking	**Contract Number**
	Grant Number
	Program Element Number
Author(s)	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Electrical and Computer Engineering Department Florida International University Miami, FL	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research, Development & Standardization Group PSC 802 Box 15 FPO AE 09499-1500	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom., The original document contains color images.	
Abstract	
Subject Terms	
Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UU
Number of Pages 4	

Using real-time image processing methods, such as, edge detection and determination of the center of gravity, the orientation of the user's gaze is assessed continuously using the relative position of these two points in the video image. This orientation, along with the necessary geometrical information that has been captured during a calibration process, prior to the use of the system, is used to compute the user's point of gaze on the computer screen.

The ASL-504 Eye Gaze tracking system is configured around an ASL-500 Controller Unit that connects to an external Control PC (different from the PC where the cursor is controlled) for setup and operation. The Control Unit is also connected to a pan-tilt module including the infrared camera and illuminators. Most importantly, the Control Unit processes de video images captured by the camera, at a rate of 60 frames per second, and determines the location of the CR and PR within the camera's field of view. This same unit uses the geometrical information gathered during the calibration of the system to convert the relative position of the CR and PR reflections into an x, y estimate of the point of gaze of the computer user on the computer screen. These estimates, along with a pupil diameter measurement and flags that indicate when a proper point of gaze estimation was not possible from the video image, are transmitted out of the Control Unit through a serial port, at a rate of 60 times per second. The different items of information are sent in a proprietary format through the serial port. ASL also provided basic C routines to read these frames of information from the serial port of the computer where the cursor control is to be implemented.

To use the ASL-504 Eye Gaze Tracking system as our EGT subsystem we developed a program that would continuously read the serial port of the computer where the cursor is controlled to receive and store the instantaneous estimates of the point of gaze, which we refer to as (POGx(n), POGy(n)). In this notation n is the discrete time index that represents the current value of the point of gaze coordinates.

B. Electromyogram (EMG) Subsystem

The EMG-based interface was originally developed at FIU in 1998, for stand-alone operation, and its design and performance are reported in detail elsewhere [1,2].

As an overview, the original goal of the EMG-based system was to command an incremental (step) cursor motion on the screen every time the system sensed the voluntary contraction of a different group of facial muscles. A cursor step UP would be commanded if the system detected that the user is raising the eyebrows. Similarly, a step DOWN would be commanded if the user is lowering the eyebrows. Clenching the right side of the jaw, by contracting the right temporalis muscle would command a cursor step to the RIGHT, while contracting the left temporalis would step the cursor LEFT.

In order to detect the activation of these muscles, the system uses only threes electrodes: E_0 in the forehead, about 3/4" to the side of the midline of the head, E_1 on the left side of the head, and E_2 on the right side of the head. All the EMG measurements were referred to an electrode placed in the left mastoid area, as illustrated in Figure 1.

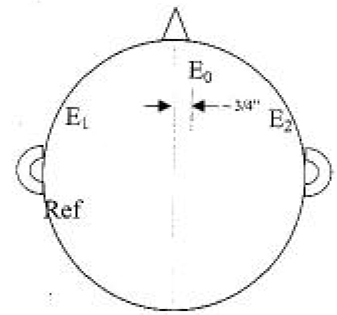

Fig.1 Placement of the electrodes used for the EMG subsystem (top view)

The simplest conceptual approach to the detection of the contraction of the different muscles monitored would rely on the identification of a temporary increase of the average power in the EMG from each electrode. However, the head of the subject acts as a volume conductor, causing "cross-talk" between the EMG signals. For example, contraction of the corrugator muscle in the forehead will still contribute a strong component in the EMG signal picked-up by E_1, which should ideally only record EMG due to the contraction of the left temporalis muscle.

To overcome the "cross-talk" problem, it was necessary to develop a detection algorithm that would discriminate EMG signals generated by the contractions of different muscles on the basis of more than just overall power. This was achieved by implementing a real-time periodogram estimation of the Power Spectral Density (PSD) of each of the three EMG signals collected, followed by a classification algorithm that focuses on the spectral differences noted in the EMG from the different facial muscles considered. The existence of spectral differences in EMG caused by different types of muscles had been reported in the past [7], and seems to be attributable to the dependence of the frequency content, specifically the mean frequency, on the contraction length of the muscle, and other factors, such as the motor unit recruitment patterns, distinct motor unit properties (fast-twitch, slow-twitch), conduction velocity, etc. [7].

Our approach was based on the observation of EMG spectra of the muscles involved. We noted that there were certain frequency bands that would have considerable differences in their relative power contents, depending on which muscles were contracted. Thus, the method would monitor the partial accumulation of power spectral density (as estimated by the periodogram method) in these bands, and establish critical comparisons, in order to identify the muscle that contracted.

The partial PSD accumulations monitored were:

F_k : From 0 Hz to 145 Hz
J_k : From 145 Hz to 600 Hz (half the sampling rate)

Where k is the electrode number considered (k = 0, 1, 2)

Based on the values of F_k and J_k found for all the channels within the processing of a given data block, the system looks for a set of conditions that will result in a muscle contraction detection, and, if one is found, the appropriate incremental cursor movement, or "click", is commanded.

For example, the conditions that the system uses to identify a left temporalis contraction and, as a consequence, command a cursor step to the left are:

Conditions for LEFT CURSOR MOVEMENT:

If	$\max(PSD_1) > Th_1$
and	$\max(PSD_0) < Th_0$, and $\max(PSD_2) < Th_2$
and	$J_1 > F_1$, and $J_1 > J_2$
Then:	LEFT cursor movement.

In these conditions, $\max(PSD_k)$ represents the largest single-bin PSD estimation obtained from the periodogram calculation on the present data block from electrode k. Th_0, Th_1 and Th_2 are pre-defined system thresholds.

The system incorporated similar sets of conditions for the detection of the right temporalis contraction and the raising and lowering of the eyebrows, which are used by the system to command cursor steps in the RIGHT, UP and DOWN directions. One additional set of conditions was used to detect the simultaneous contraction of both temporalis muscles (full jaw clench), which is used by the system to command a mouse "click" operation. These sets of conditions can be found in previous reports focusing on the EMG-based system alone [1,2].

According to the preceding explanation, the EMG-subsystem generates an output every time a data block is processed, which is approximately four times per second. The output may be a command to step the cursor, UP, DOWN, LEFT or RIGHT, a command to perform a click, or it may be a NULL output, if none of the sets of conditions was fulfilled. In terms of cursor movement control exclusively, the output of this EMG subsystem could be represented by two variables: $\Delta x(n)$, $\Delta y(n)$, which can only take on values of 1, -1 and 0.

All the processing involved in the implementation of the EMG subsystem takes place in a plug-in A-to-D/DSP board (ADC64, by Innovative Integration), installed in the computer whose cursor is being controlled. So, the results from this subsystem are directly available to the program that was developed to merge the outputs of the two subsystems (EGT and EMG) according to the assessment of the context in which the user is providing his/her input to the computer.

C. Context Assessment and effective control definition

The previous two sections have outlined the architecture and functionality of both the EGT- and the EMG- based cursor control modules. The functionality of each of these approaches to cursor control reveals their complementary strengths. While the EGT-based cursor control approach excels in providing fast, broad displacements of the cursor, over long distances across the computer screen, this method presents significant shortcomings when steady, small displacements are needed, such as the ones required to place the cursor on small GUI icons. Similarly, previous studies [5] have reported the limitations and inaccuracies of EGT-based "click protocols", like the "wink" and "dwell" protocols. On the other hand, the EMG-based cursor control developed at FIU showed high stability for small displacements, and the ability to reliably detect the full-jaw clench to instruct a "click". However, long cursor excursions were slow with this interface, given its stepping nature. For the integration of these two modalities in a single hybrid EGT/EMG cursor control approach, we devised a context detection scheme based on the estimated intent of the user.

We estimated that while the user is involved with a small neighborhood of the screen, around the current position of the cursor, (Cx(n),Cy(n)) the EMG-based control should be enabled and the EGT-control should be disabled. In this way, the cursor will remain static unless the user performs muscle contractions to command short, steady displacements or a "click" operation. On the other hand, if the user intends to interact with areas of the screen that are at a considerable distance from the current position of the cursor, he/she will first direct his/her gaze to that section of the screen. It is at this moment that the intent of the user should be detected to switch the effective control of the cursor to EGT-based control, which can quickly re-position the cursor at the current location of the user's point of gaze. For this context-switching approach the key measurement is what we called "POG drive", which is the instantaneous distance between the estimated POG and the previous cursor position:

$$POG\_drive = \sqrt{\left(POGx(n) - Cx(n-1)\right)^2 + \left(POGy(n) - Cy(n-1)\right)^2} \quad (1)$$

The context detection algorithm evaluates this POG_drive value every time a new set of values is read from the EMG submodule (i.e. 4 times per second). If the POG_drive is found to be larger that a preset Radius Threshold, *R*, (e.g. 80 pixels), the algorithm assumes that the user is driving his/her gaze away from the current cursor position and it enables the EGT-based cursor control:

$$Cx(n) = POGx(n) \quad (2)$$

$$Cy(n) = POGy(n) \quad (3)$$

If, on the other hand, it is found that the POG_drive is less than the Radius Threshold, R, then the EMG-based control is enabled and the resulting cursor position will be modified according to the values passed by the EMG submodule:

$$Cx(n) = Cx(n-1) + \Delta x(n) \quad (4)$$

$$Cy(n) = Cy(n-1) + \Delta y(n) \quad (5)$$

III. RESULTS

The efficiency of the hybrid EMG/EGT interface was tested through an evaluation protocol that exercises the pointing and clicking abilities of the interface. This protocol presents a Start Button to the user in one corner of the screen. The dimensions of the Start Button are always 8.5 x 8.5 mm. The protocol also shows the user a Stop Button, always at the center of the screen. There are four sizes for this target: 8.5 x 8.5 mm; 12.5 x 12.5 mm; 17 x 17 mm; and 22 x 22 mm. Before the beginning of each trial the cursor is placed for the user at the Start Button. Then the subject is to use the hybrid system to a) Click on the Start Button, to start a timer, b) Move the cursor towards the Stop Button, following any trajectory, and c) Click on the Stop Button, to stop the timer. At the end of each trial the time, in seconds, taken by the user for the trial is displayed. These timing results are logged in a text file, for further analysis.

Each test session consisted of 20 trials with each size of Stop Button, which varied from smallest to largest, for a total of 80 trials. Within each group of 20 trials with the same Stop Button size the Start Button position was rotated through the four corners of the screen, from one trial to the next. So, there were 5 trials starting at each corner for each Stop Button size. Six college-aged subjects participated in the evaluation. In our previous development of the EMG-based interface, we had used the same evaluation protocol to assess the efficiency of the EMG-only interface, as well as a normal hand-held mouse interface. Figures 2 and 3 summarize the results of the evaluation, where the data has been consolidated by icon size, across all 6 experimental subjects, for the 3 interfaces tested.

IV. DISCUSSION

Figure 2 shows that the hybrid EGT/EMG interface achieves an appreciable reduction in the time needed to complete the evaluation task, with respect to the EMG-only interface. Furthermore, for all but the smallest icon size, the performance does not seem to be as closely correlated to icon size as it was for the EMG-only interface. The average times for the standard hand-held mouse interface are, of course, much smaller, and are shown just for reference. Similarly, the hand-held mouse displayed minimum variability in Figure 3, while the hybrid system recorded the largest variability of all. We believe that this apparently larger inconsistency in the performance with the hybrid system is due to the increased complexity of this bi-modal interface. It seems that learning how to skillfully take advantage of the two forms of interaction offered by the system may require some training period, which was not allowed in the experiments reported here.

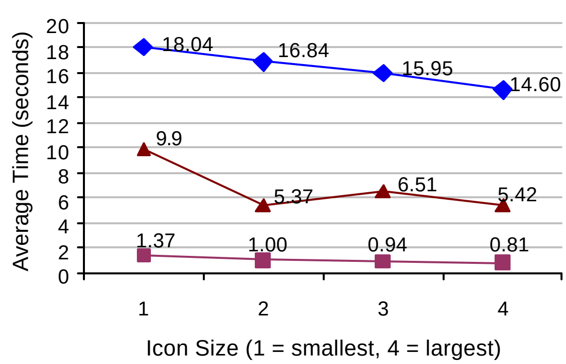

Fig.2. Average trial times, by Stop Button size, for the three interfaces considered: Hand-held mouse (squares); EMG-only system (diamonds); Hybrid system (triangles)

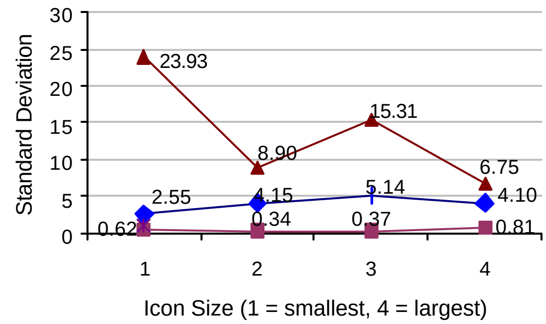

Fig. 3. Standard Deviation for trial times, by Stop Button size, for the three interfaces considered: Hand-held mouse (squares); EMG-only system (diamonds); Hybrid system (triangles)

V. CONCLUSION

The hybrid EGT/EMG human-computer interface presented in the paper combines the complementary strengths of the EGT-only systems (quick broad displacements of the cursor) and our EMG-only interface (steady cursor, accurate small displacements, reliable click). In the point-and-click experiment used to evaluate its efficiency, the hybrid interface showed a reduction of approximately 50% of the time required for the task, with respect to the EMG-only interface. Future experiments may confirm the need for a "training period" for the efficient use of this interface.

REFERENCES

[1] Barreto, A. B., Scargle, S. D., and Adjouadi, M., "Hands-off Human-Computer Interface for Individuals with Severe Motor Disabilities", in Human-Computer Interaction: Communication, Cooperation and Application Design, Vol. 2, pp. 970-974. Bullinger, H-J., and Ziegler, J.(Eds.), Lawrence Erlbaum Associates, 1999.

[2] Barreto, A. B., Scargle, S. D., and Adjouadi, M., "A Practical EMG-based Human-Computer Interface for Users with Motor Disabilities", Journal of Rehabilitation Research & Development, Vol. 37(1), pp. 53-63, 2000.

[3] Chi C-F, and Lin C-L., Speed and Accuracy of Eye-Gaze Pointing", Perceptual and Motor Skills, 85, 705-718, 1997.

[4] Hutchinson T., White K., Martin W., Reichert K., and Frey L., "Human-Computer Interaction Using Eye-Gaze Input", IEEE Transactions on Systems, Man, and Cybernetics, Vol., 19, No. 6, pp. 1527 – 1534, 1989.

[5] Istance H., and Howarth P., "Keeping an eye on your interface: The Potential for eye-based control of Graphical User Interfaces (G.U.I.'s)", Proceedings of the HCI International '94, The International Conference on Human-Computer Interaction, 1994.

[6] Jacob R., "The Use of Eye Movements in Human-Computer Interaction Techniques: What you Look At is What You Get". ACM Transactions on Information Systems, Vol. 9, No. 3, pp. 152 – 169, April 1991.

[7] LeVeau B. and Andersson G., "Output Forms: Data Analysis and Applications. Interpretation of the Electromyographic Signal", Dept. of Health and Human Services, NIOSH Pub. No. 91-100, 1992.